

Article



# Integrative Proteome and Transcriptome Analyses Reveal the Metabolic Disturbance of the Articular Cartilage in Kashin–Beck Disease, an Endemic Arthritis

Lixin Han<sup>+</sup>, Bolun Cheng<sup>\*,+</sup>, Jinyu Xia, Shiqiang Cheng, Xuena Yang and Feng Zhang<sup>\*</sup>

NHC Key Laboratory of Environment and Endemic Diseases, Collaborative Innovation Center of Endemic Disease and Health Promotion for Silk Road Region, School of Public Health, Health Science Center, Xi'an Jiaotong University, Xi'an 710061, China; shaanxiyqfk@163.com (L.H.); xjy0626@stu.xjtu.edu.cn (J.X.); chengsq0701@stu.xjtu.edu.cn (S.C.); smile940323@stu.xjtu.edu.cn (X.Y.)

\* Correspondence: boluncheng@xjtu.edu.cn (B.C.); fzhxjtu@mail.xjtu.edu.cn (F.Z.)

<sup>+</sup> These authors contributed equally to this work.

**Abstract:** The objective of this study was to elucidate the proteomic and transcriptomic alterations within the cartilage in Kashin-Beck disease (KBD) compared to a normal control. We conducted a comparison of the expression profiles of proteins, mRNAs, and lncRNAs via data-independent acquisition (DIA) proteomics and transcriptome sequencing in six KBD individuals and six normal individuals. To facilitate the functional annotation enrichment analysis of the differentially expressed (DE) proteins, DE mRNAs, and DE lncRNAs, we employed bioinformatic analysis utilizing Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG). Additionally, we conducted integration analysis of multiomics datasets using mixOmics. We revealed a distinct proteomic signature, highlighting 53 DE proteins, with notable alterations in the pathways related to tryptophan metabolism and microbial metabolism. Additionally, we identified 160 DE mRNAs, with the functional enrichment analysis uncovering pathways related to RNA metabolism and protein splicing. Furthermore, our analysis of the lncRNAs demonstrated biological processes involved in protein metabolism and cellular nitrogen compound metabolic processes. The integrative analysis uncovered significant correlations, including the positive correlation between superoxide dismutase 1 (SOD1) and mitochondrial import receptor subunit TOM6 homolog (TOMM6), and the negative correlation between C-X9-C motif-containing 1 (CMC1) and succinate-CoA ligase [GDP-forming] subunit beta, mitochondrial (SUCLG2). Our results provide novel insights into the molecular mechanisms underlying KBD.

Keywords: Kashin–Beck disease; osteoarthritis; proteome; transcriptome; multi-omics

# 1. Introduction

Kashin–Beck disease (KBD) is a chronic, endemic osteochondropathy that primarily affects children and adolescents. This disease is characterized by significant pathological changes in the growth plate and articular cartilage, leading to necrosis of the chondrocytes, which disrupts normal endochondral ossification [1]. Consequently, this disruption results in a range of clinical manifestations, including growth retardation, joint deformities, and functional impairments of multiple joints [2,3]. Typically, the disease presents in children aged 3 to 12 years, with the symptoms becoming more pronounced as they age [4]. In advanced stages, KBD can lead to severe complications, including dwarfism and permanent disability, significantly impacting quality of life [5]. The pathogenesis of KBD is closely



Academic Editor: Nikolay A. Kolchanov

Received: 8 April 2025 Revised: 22 May 2025 Accepted: 26 May 2025 Published: 27 May 2025

Citation: Han, L.; Cheng, B.; Xia, J.; Cheng, S.; Yang, X.; Zhang, F. Integrative Proteome and Transcriptome Analyses Reveal the Metabolic Disturbance of the Articular Cartilage in Kashin–Beck Disease, an Endemic Arthritis. *Int. J. Mol. Sci.* **2025**, *26*, 5146. https://doi.org/ 10.3390/ijms26115146

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/ licenses/by/4.0/). linked to genetic alterations, making the investigation of key genes and proteins essential for understanding the disease's progression and for developing potential therapeutic strategies.

Recent research on KBD has increasingly focused on transcriptomic analyses to elucidate the underlying biological mechanisms, particularly in relation to osteoarthritis (OA). For instance, RNA sequencing has highlighted differences in the gene expression profiles between KBD with OA and rheumatoid arthritis (RA), identifying key pathways involved in cartilage metabolism and degradation [6–9]. While mRNAs are traditionally the focus of transcriptomic studies, lncRNAs are increasingly recognized as pivotal regulators of gene expression at the transcriptional, post-transcriptional, and epigenetic levels [10]. Emerging evidence supports lncRNAs playing a key role in cartilage degradation and skeletal diseases [11]. However, a notable limitation of the existing research is the lack of normal controls, which limits comprehensive understanding of the disease mechanisms. The integration of transcriptomic and proteomic data remains insufficient, underscoring the necessity of combining multi-omics analyses to explore the mechanisms of KBD.

DIA proteomics has emerged as a powerful technique in the field of large-scale proteomics, offering significant advantages over traditional data-dependent acquisition (DDA) methods [12,13]. The high reproducibility and sensitivity of DIA make it particularly suitable for analyzing small amounts of clinical tissue specimens [14]. For bone-related diseases, DIA has been effectively applied to investigate the proteomic landscape of the cartilage and subchondral bone, providing insights into disease pathogenesis. For instance, a study utilizing DIA mapped the cartilage and subchondral bone proteomes, identifying multiple key proteins involved in osteogenesis and bone remodeling [15]. However, comprehensive studies utilizing DIA to compare the proteomes of KBD and normal control cartilage tissues remain lacking.

In this study, we employed DIA proteomics and RNA sequencing (RNA-seq) to profile the proteomic and transcriptomic alterations in KBD by comparing the joint cartilage of KBD patients to that of normal controls. Our integrated analysis revealed DE mRNAs, lncRNAs, and proteins, elucidating their functional implications. These findings enhance our molecular understanding of KBD and provide a foundation for developing targeted therapeutic strategies based on these biomarkers.

## 2. Results

#### 2.1. Proteomic Profiles of Cartilage from Patients with and Without KBD

Our comparative proteomic analysis revealed a distinct proteomic signature between KBD and normal cartilage (Figure 1A), highlighting 53 proteins with significant differential expression (Supplementary Table S1). Among these, 18 proteins were upregulated, while 35 were downregulated (Figure 1B). For example, ATP8B1, implicated in phospholipid transport; A1BG, known for its acute-phase protein properties; SUCLG2, a key enzyme in the citric acid cycle; ANXA1, an anti-inflammatory protein; and TOMM6, a component of the mitochondrial import machinery. Three pathways emerged as significantly altered: tryptophan metabolism, microbial metabolism in diverse environments, and biosynthesis of secondary metabolites (Supplementary Table S2). Furthermore, our analysis identified 15 significant GO terms under the cellular component (CC) (Figure 1C), including organelle membrane contact sites and cytoplasmic vesicles, suggesting alterations in intracellular trafficking and organelle interactions. Under the biological process (BP) (Figure 1D), 117 significant GO terms were noted, such as chromosome organization, DNA conformation changes, and protein–DNA complex assembly, which may suggest disruptions to cellular processes and genetic regulation. For the molecular function (MF) (Figure 1E), 19 significant GO terms were highlighted, focusing on calcium-dependent phospholipid



binding and various enzymatic activities, reflecting changes in signal transduction and metabolic regulation (Supplementary Table S2).

**Figure 1.** Expression profiles of proteins. Overview of the protein expression (KBD n = 6, control n = 6). (**A**) The heatmap is used to assess the expression of the proteins. Each DE protein is represented by a single row of colored boxes, and each sample is represented by a single column. (**B**) Volcano plots of the proteins. Red and blue denote high and low expression, respectively (p value < 0.05 and  $|\log_2 FC| > 1$ ). (**C**–**E**) CC, BP, and MF GO term analysis of the DE proteins (Fisher's exact test p value < 0.05).

#### 2.2. mRNA Differences in Cartilage from Patients with and Without KBD

Our study revealed significant transcriptional differences between the KBD and control samples (Figure 2A), with 158 DE mRNAs identified (Supplementary Table S3), of which 30 were upregulated and 128 were downregulated in KBD (Figure 2B). The functional enrichment analysis of these DE mRNAs uncovered 226 significant pathways (Figure 2C, Supplementary Table S4), including the metabolism of RNA, metabolism of proteins, mRNA splicing—major pathway, and infectious disease. Furthermore, 316 significant GO terms under BP were enriched, highlighting processes such as mRNA metabolic process, RNA splicing, and cellular macromolecule catabolic process. Under CC, 79 significant GO terms were identified, including the intracellular organelle part and cytoplasm. For MF, 32 significant GO terms were enriched, focusing on protein binding, RNA binding, and nucleic acid binding (Figure 2D, Supplementary Table S4). To assess the transcriptomic–proteomic concordance, we examined the mRNA fragments per kilobase of transcript per million

mapped reads (FPKM) of genes encoding the top five DE proteins. Although these mRNAs did not reach statistical significance, directional trends were observed. For instance, both *TMSB10* and *ANXA1* exhibited upregulation in the KBD samples at both the RNA and protein levels (Supplementary Figure S1).



**Figure 2.** Expression profiles of mRNAs. Overview of the mRNA expression (KBD n = 3, control n = 3). (**A**) The heatmap is used to assess the expression of the mRNAs. Each DE mRNA is represented by a single row of colored boxes, and each sample is represented by a single column. (**B**) Volcano plots of the mRNAs. Red and blue denote high and low expression, respectively (p value < 0.05 and  $|\log_2 FC| > 1$ ). (**C**) Pathway analysis of the DE mRNAs (Fisher's exact test p value < 0.05). (**D**) GO term analysis of the DE mRNAs (Fisher's exact test p value < 0.05).

## 2.3. Differential IncRNA Expression and Target Gene Prediction and Functional Analysis

Our analysis revealed a significant alteration in the lncRNA landscape between the KBD and control samples (Figure 3A), with 218 lncRNAs exhibiting differential expression— 111 being upregulated and 107 downregulated in KBD (Figure 3B, Supplementary Table S5). To identify potential functional targets of the DE lncRNAs, we employed BLAST2GO to predict the target genes through sequence homology analysis. We identified 60 lncRNAs that target DE genes (Supplementary Table S6). For the predicted target genes of the DE lncRNAs, a total of 192 significant pathways were enriched (Figure 3C, Supplementary Table S7), such as the metabolism of proteins, mRNA splicing, and signaling by Wnt. Under the GO categories, we identified 170 significant BP terms, including RNA splicing and cellular nitrogen compound metabolic process. For CC, 79 significant terms were enriched, such as the intracellular organelle part and cytoplasm. For MF, 19 significant terms were identified, including protein binding and RNA binding (Figure 3D, Supplementary Table S7).



**Figure 3.** Expression profiles of long intergenic ncRNAs. (**A**) The heatmap is used to assess the expression of the lncRNAs. (**B**) Volcano plots of the lncRNAs. (**C**) Pathway analysis of the target genes of the DE lncRNAs. (**D**) GO analysis of the target genes of the DE lncRNAs.

## 2.4. Functional Enrichment Comparison Among Multi-Omics Data

Due to the lack of overlap between the RNA and protein data, we investigated whether there was overlap at the pathway or ontology level. The tri-omics (lncRNA–mRNA– protein) pathway and ontology overlap included two GO terms, metabolic process and cellular macromolecular complex assembly. These GO terms reflect core cellular processes critical for maintaining cartilage homeostasis. The mRNA–protein overlap included eight GO terms, such as secretion by cell, cell cycle G2/M transition, and RNA biosynthetic process. The lncRNA–protein overlap included one GO term, single-stranded RNA binding. lncRNAs may regulate protein activity by binding RNA-binding proteins. The lncRNA– mRNA overlap included 135 GO terms, such as immune/inflammatory regulation, RNA metabolic process, intracellular organelle, and regulation of macromolecule metabolic process (Supplementary Table S8). The lncRNA–mRNA overlap included 123 pathways, such as the metabolism of proteins, regulation of apoptosis, and MAPK family signaling cascades (Supplementary Table S9). The extensive lncRNA–mRNA overlap highlights the transcriptional and post-transcriptional regulatory networks in KBD.

## 2.5. Integrative Analysis of Proteome and Transcriptome

Figure 4A shows the correlation of the top 20 features in the three datasets. The first component of the sparse partial least squares discriminant analysis (sPLS-DA) [16] of the combined proteome and transcriptome datasets clearly discriminated the normal from the KBD cartilage samples (Figure 4B). This filtering identified proteins, mRNAs, and IncRNAs that distinguished KBD from the normal control samples (Figure 4C). For example, Q9BRF8 (serine/threonine-protein phosphatase CPPED1), P04217 (alpha-1B-glycoprotein), Q86X76 (deaminated glutathione amidase), Q96B49 (mitochondrial import receptor subunit TOM6 homolog), and P68371 (tubulin beta-4B chain) were highlighted. For example, P04217 (alpha-1B-glycoprotein) and NONHSAG048636.2 were KBD-specific, which may potentially distinguish the KBD samples from the controls. The DIABLO method identified several important features discriminating KBD from the normal controls by interrogating the correlations between the three omics datasets, with the lncRNA and proteome data showing the highest correlations (Pearson's r = 0.96) (Figure 4D). Finally, to visualize the between-omics correlations in the DIABLO analysis, a Circos plot revealed numerous strong positive and negative correlations (Pearson's r > 0.70) (Figure 4E). For example, SOD1 was positively correlated with Q96B49 (TOMM6). CMC1 was negatively correlated with Q96I99 (SUCLG2) and positively correlated with P68371 (tubulin beta-4B chain).



**Figure 4.** Sparse partial least squares discriminant analysis. (**A**) Correlation circle plot for the combination of the three datasets. The variable types are indicated with different symbols and colors, overlaid on the same plot. (**B**) The individual contribution of each dataset, showing the score plots for the first two components, indicating the efficient separation capability for the three datasets. The samples are plotted according to their scores for the first two components for each dataset, colored by KBD or control. (**C**) Top 5 selected features shown in the pyramid bar plot. The most important variables (according to the absolute value of their coefficients) are ordered from bottom to top. As this is a supervised analysis, the colors indicate the class for which the median expression value is the highest for each feature. The X-axis represents the correlation coefficient (r) between the molecule and KBD or the control. (**D**) Sample scatterplot from plotDiablo displaying the first component in each dataset (upper diagonal plot) and the Pearson's correlation between each component (lower diagonal plot). The samples are colored by KBD or control, with the 95% confidence ellipse plots represented. (**E**) The Circos plot (cutoff: 0.7) shows positive or negative correlations greater than 0.7 between variables of different types, denoted as red and blue lines, respectively.

# 3. Discussion

In this study, we established the proteomic and transcriptomic profiles of the articular cartilage in KBD and normal control samples, aiming to elucidate the molecular mechanisms of KBD. Our comprehensive analysis included proteins, lncRNAs and mRNAs, and

it identified complex network of features and gene–protein interactions, offering novel insights into KBD's etiology.

Our findings identified metabolic and oxidative stress pathways as central to KBD's pathogenesis. For example, low selenium levels and T-2 toxin exposure have been implicated in cartilage damage through mechanisms involving mitochondrial dysfunction and reactive oxygen species (ROS) accumulation [17]. Our multi-omics data corroborated this and revealed that downregulation of SUCLG2 (succinate-CoA ligase) and upregulation of SOD1 align with the mitochondrial energy metabolism defects and compensatory antioxidant responses observed in KBD chondrocytes [18]. In addition, the enrichment of the microbial metabolism pathways suggests a potential link between gut microbiota dysbiosis and KBD progression, a hypothesis supported by a recent genomic study [19]. However, our study provides novel insights by integrating proteomic and transcriptomic data to identify the lncRNA-mediated regulation of protein metabolism. While both KBD and OA involve cartilage degradation, our findings highlight key molecular distinctions. OA is primarily driven by mechanical stress and age-related inflammation (e.g., IL-1 $\beta$ , TNF- $\alpha$ ) [20], whereas KBD exhibits early-onset mitochondrial dysfunction and a selenium-dependent redox imbalance [21]. In OA, lncRNAs such as plasmacytoma variant translocation 1 (PVT1) regulate chondrocyte apoptosis and matrix metalloproteinase (MMP) expression [22]. In contrast, our study identifies KBD-specific lncRNAs enriched in nitrogen compound metabolism.

The transcriptome and proteomics analysis of KBD found that there were no overlapping molecules. This finding indicated the potential roles of post-transcriptional regulation and epigenetic regulation in this process. Firstly, the absence of overlapping genes between the transcriptome and the proteome is not uncommon. Research has indicated that the regulation of gene expression is influenced not only by the transcription level but also by various factors, including post-transcriptional modification, post-translational modification and protein degradation. For example, numerous genes may be regulated by lncRNAs and other mechanisms after transcription [23]. Secondly, epigenetic alterations may contribute to the inconsistencies observed between the transcriptome and the proteome. Epigenetic modifications such as DNA methylation and histone modification can regulate gene expression without altering the DNA sequence. These alterations may be crucial to the pathogenesis of KBD, as certain epigenetic modifications have been strongly linked to the onset and progression of arthritis [24]. Additionally, the intricate environmental risk factors associated with KBD may complicate the identification of key genes. Variations in these processes can influence gene expression patterns, resulting in differences between the transcriptomic and proteomic profiles.

Our integrated omics analysis has revealed a range of proteins and transcripts related to metabolism that could serve as potential biomarkers for KBD. For instance, ANXA1, a protein known for its anti-inflammatory and immunomodulatory functions, has been associated with the resolution of inflammation in arthritic conditions [25]. TMSB10, recognized for its contributions to angiogenesis and osteogenesis, has been applied in bone tissue engineering to enhance bone regeneration [26]. SUCLG2, which is part of the citric acid cycle, has been shown to be upregulated by fibroblast growth factor 2 (*FGF2*) in osteoblasts, highlighting its significance in relation to bone development and metabolism [27]. The differential expression of these proteins in KBD implies that they may have crucial impacts on cartilage metabolism. Upon differential gene expression analysis, we noticed many DEGs were actually lncRNAs and therefore decided to investigate the target genes of these lncRNAs. An RNA-seq study identified 4103 DE lncRNAs in the KBD cartilage, implicating pathways such as Wnt signaling, lysosome function, and TNF signaling [11]. These pathways align with our proteomic findings, suggesting the lncRNAs may bridge the transcriptional and metabolic dysregulation in KBD. In OA, the lncRNAs regulate chon-

drocyte apoptosis and matrix degradation [28]. Similarly, our study identified lncRNAs enriched in protein metabolism and nitrogen compound processes, potentially suggesting these regulatory roles in KBD.

Our enrichment analyses revealed that several functions within bone- and cartilagerelated pathways may be critical to KBD's pathogenesis. The enrichment analysis of the DE proteins underscored the potential metabolic dysregulation associated with KBD, particularly concerning tryptophan metabolism, which has been linked to inflammatory responses [29]. Moreover, the identified GO biological processes suggested that changes in genomic stability and transcriptional regulation may play a role in the disease's progression [30]. The enrichment analysis of the DE mRNAs highlighted pathways related to RNA metabolism and splicing, as well as immune system functions. The involvement of mRNA splicing pathways emphasized the significance of post-transcriptional regulation in KBD, which may influence the expression of key inflammatory mediators [31]. Additionally, the GO biological processes associated with cellular nitrogen compound metabolism further highlighted the metabolic alterations occurring in KBD, potentially affecting cellular homeostasis and immune responses. The identification of RNA splicing and intracellular transport processes among the enriched GO terms for the lncRNA targets suggested that the lncRNAs may modulate the expression and function of proteins involved in these essential pathways. The tri-omics overlaps highlighted metabolic pathways vulnerable to KBD stressor selenium deficiency. The tri-omics enrichment of the metabolic pathways supports the hypothesis that mitochondrial failure underlies KBD chondrocyte degeneration. In addition, the immune pathways, such as NF-KB, overlapped in the lncRNAs and mRNAs suggested inflammation progress in KBD. Collectively, these findings indicate that immune responses, alterations in the endocrine system, and metabolic changes represent the primary functional shifts observed between KBD and controls.

The integration of data through multi-omics approaches is essential for studies on bone and cartilage metabolism [32]. Our integrative analysis has provided a comprehensive perspective on the molecular alterations in KBD patients, emphasizing the association between the proteome and the transcriptome in regulating phenotypes [33]. We identified numerous proteins, mRNAs, and lncRNAs that differentiate KBD from normal controls. The *SOD1* gene, a crucial component of the oxidative stress response, showed positive correlations with TOMM6, a subunit of the mitochondrial import receptor. Previous studies have indicated that *SOD1* is vital to managing oxidative stress [34,35] and maintaining bone homeostasis [36]. The upregulation of *SOD1* and downregulation of glutathione-related enzymes suggested that boosting the endogenous antioxidant defenses (e.g., N-acetylcysteine [37], selenium-rich diets [38]) may mitigate the oxidative damage in KBD cartilage. Moreover, the microbial metabolism pathway enrichment suggested that probiotics or dietary modifications to reduce T-2 toxin exposure may disrupt KBD progression, aligning with the fungal toxin hypothesis of KBD's etiology [39].

Our research utilized high-throughput technologies to characterize KBD from various perspectives, encompassing both proteomic and transcriptomic analyses. DIABLO provides a robust statistical framework for integrating disparate datasets—comprising RNA and protein data. This approach offers a comprehensive view of the molecular landscape of KBD. However, there are several limitations and future directions. Firstly, the modest sample size may reduce the statistical power to detect subtle transcriptomic and proteomic differences. While our multi-omics integration approach partially mitigates this issue by prioritizing pathways across datasets, future studies with larger cohorts are needed to validate these findings. Secondly, while this study focused on multi-omics discovery, functional validation of the prioritized targets using in vitro or in vivo models is essential to confirm their causal roles in KBD's pathogenesis. Thirdly, it is difficult to correlate

molecular signatures with clinical outcomes due to the small cohort size and limited clinical metadata. Future studies integrating multi-omics profiling with detailed phenotyping in larger KBD cohorts are essential.

In summary, our study presents a comprehensive analysis of the transcriptional and proteomic alterations in KBD, highlighting the potential of integrated omics approaches in unraveling the complexities of the disease pathology. By examining the functional pathways and differential expression of RNAs and proteins, we identified key genes and non-coding RNAs that may provide novel insights into KBD's etiology.

# 4. Materials and Methods

## 4.1. Ethics Statement

This study was approved by the Human Ethics Committees of Xi'an Jiaotong University. Written informed consent was obtained from all the study participants.

#### 4.2. Samples

Cartilage specimens were obtained from the knee joints of patients diagnosed with KBD and from age- and sex-matched control subjects. The diagnosis of KBD was established based on clinical evaluations and imaging studies, adhering to the KBD diagnosis criterion of China (WS/T 207-2010 [40], https://www.ndcpa.gov.cn/jbkzzx/c100200/common/content/content\_1666632341665943552.html accessed on 25 May 2025). Participants with a history of genetic bone and cartilage disorders, RA, or other significant joint diseases were excluded from this study. The KBD group included six individuals (three males and three females) aged between 54 and 75 years (Figure 5), while the control group comprised six subjects (three males and three females) aged between 57 and 71 years. For the proteomic analysis, a total of six KBD and six control samples underwent differential protein expression analysis using DIA liquid chromatography–mass spectrometry (LC-MS). Additionally, RNA sequencing was performed on three KBD samples and three control samples to investigate the gene expression profiles. The average ages of the KBD and control groups were  $66.5 \pm 7.21$  and  $63.5 \pm 5.12$  years, respectively.



**Figure 5.** Characteristics of the study samples with Kashin–Beck disease. (**A**) Hand images of male patient; (**B**) hand images of female patient; (**C**) hand X-ray of male patient; and (**D**) hand X-ray of female patient. The image (**A**) and X-ray (**C**) are from the same patient, while the image (**B**) and X-ray (**D**) are from the same patient.

#### 4.3. Proteomic Analysis

#### 4.3.1. Sample Preparation

For the proteomic analysis, cartilage samples from KBD patients and control subjects were processed following a standardized protocol. Initially, proteins were extracted from the cartilage tissues using a lysis buffer containing 8 M urea and 50 mM ammonium bicarbonate. The protein concentration was determined using the Bradford assay, ensuring accurate quantification prior to further processing. Subsequently, the extracted proteins underwent enzymatic digestion with trypsin at 37 °C for 16 h, allowing for the generation of peptides suitable for mass spectrometry analysis. Following digestion, the samples were desalted using C18 solid-phase extraction to remove any contaminants and concentrate the peptides. Finally, the purified peptides were reconstituted in a suitable buffer for library construction and mass spectrometry analysis.

#### 4.3.2. LC-MS Analysis

The prepared peptide samples were subjected to LC-MS (Shimadzu, Inc., Shanghai, China) for comprehensive proteomic profiling. The samples were loaded onto a C18 column and separated using a gradient elution method. Mass spectrometry was performed on a Q Exactive mass spectrometer (ThermoFisher, Inc., Shanghai, China), operating in both DDA and DIA modes. The DDA mode was utilized for spectral library generation, while the DIA mode allowed for the quantification of proteins across the samples. The mass spectrometer settings included a resolution of 70,000 for the MS1 scans and 17,500 for the MS2 scans, with a scan range of 300 to 1500 m/z. The data generated were processed to identify and quantify proteins based on their peptide sequences.

#### 4.3.3. DIA Proteomics Analysis

To facilitate the analysis of the mass spectrometry data, a customized spectral library was constructed. The raw data files obtained from the LC-MS analysis were processed using MaxQuant (v2.0.0), which enabled the identification of peptides and proteins based on the UniProt protein database. The parameters for the database searching included trypsin digestion with a maximum of two missed cleavages and a mass tolerance of 20 ppm for precursor ions. The false discovery rate (FDR) for the protein identification was set at 1%. The searching result was exported in a .tsv file format containing the annotation of the precursors and fragment ions and their exact retention times. The .tsv file was then imported into DIA-NN (v1.8.1) to generate the spectral library used for the DIA data analysis. The DIA data were analyzed with DIA-NN, a mass spectrometer vendor-independent free software for DIA data analysis. The raw data were analyzed according to the user guide for the software, with default settings established for the protein identification and peak area calculation. A target-decoy-based strategy was applied to control the FDR at lower than 1%. A *p* value < 0.05 and  $|\log_2FC| > 1$  were set as cutoffs to identify significantly dysregulated proteins.

#### 4.3.4. Functional Enrichment Analysis

The biological significance of the DE proteins was assessed through enrichment analysis. GO enrichment analysis was performed to categorize the proteins based on their BP, MF, and CC, implemented using the GOseq (v3.21) [41]. Additionally, KEGG pathway analysis was conducted to identify the pathways associated with the DE proteins, utilizing KOBAS (v3.0) [42]. The enrichment was considered significant for GO terms and KEGG pathways with a corrected *p* value of less than 0.05.

#### 4.4. Transcriptomic Analysis

# 4.4.1. RNA Extraction, Library Preparation, and RNA-Sequencing

The total RNA was isolated and purified using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer's procedure. A targeted approach for RNA extraction was employed, focusing on the removal of ribosomal RNA (rRNA) to enrich the mRNA and long non-coding RNA (lncRNA). This process was facilitated using the Ribo-Zero<sup>TM</sup> rRNA Removal Kit (Illumina, San Diego, CA, USA). The subsequent library preparation was carried out following the manufacturer's protocol, and the library was sequenced on the Illumina HiSeq 4000 platform (Illumina, Inc., San Diego, CA, USA), utilizing paired-end sequencing technology. The post-sequencing data processing involved a series of bioinformatics tools, including Cutadapt (v5.0) [43] for adapter trimming and quality control, FastQ (v0.14.0) for sequence quality assessment (http: //www.bioinformatics.babraham.ac.uk/projects/fastqc/, accessed on 10 January 2025), Bowtie2 (v2.5.2) [44] for aligning the reads to the reference genome, Hisat2 (v2.1.0) [45] for mapping the reads to the genome of hg38, StringTie (v1.3.3) [46] for transcript assembly and quantification, and edgeR (v3.18.1) [47] for estimating the expression levels of all the transcripts.

#### 4.4.2. LncRNA Identification

Transcripts that overlapped with known mRNAs and those shorter than 200 bp were discarded. We utilized CPC (v0.9r2) [48] and CNCI (v2.0) [49] to predict transcripts with coding potential. All the transcripts with a CPC score of <-1 and a CNCI score of <0 were removed, and the remaining transcripts were considered lncRNAs. We focused on long intergenic non-coding RNAs, defined as transcripts of >200 nucleotides lacking protein-coding potential, to investigate their regulatory roles in KBD.

#### 4.4.3. Different Expression Analysis of mRNAs and lncRNAs

StringTie [46] was used to perform the expression level calculations for the mRNAs and lncRNAs by calculating the FPKM [50]. The DE mRNAs and lncRNAs were selected with  $|\log_2 FC| > 1$ , with statistical significance defined as an adjusted *p* value < 0.05, using the R package edgeR (v3.18.1) [47].

## 4.4.4. Target Gene Prediction and Functional Analysis of lncRNAs

To elucidate the functional implications of the identified lncRNAs, we employed BLAST2GO (v6.0.1) [51] for the prediction of potential target genes. This tool enables the functional annotation of genes based on the sequence homology, providing insights into the potential biological roles of lncRNAs. By focusing on genes in proximity to lncRNAs, we aimed to uncover the cis-regulatory mechanisms that may influence the gene expression patterns. Significant differences in gene expression were assessed with  $|\log_2 FC| > 1$  and p < 0.05, followed by GO and pathway analysis. Enrichment analysis was performed to calculate the *p* value by Fisher's exact test, with p < 0.05 set as the significance threshold.

#### 4.5. Integration Analysis of Multi-Omics Datasets

To uncover the intricate relationships and novel insights that span different molecular levels, we conducted an integrated analysis using mixOmics (v6.10.9) [52], which facilitates the integration of heterogeneous omics data through multivariate statistical approaches. Our analysis focused on 3 KBD and 3 healthy controls. From each of the three omics layers—proteomics, mRNA, and lncRNA—we selected the top 20 features based on their statistical significance. The integration process leveraged the Data Integration Analysis

for Biomarker discovery using DIABLO framework [53], which is embedded within the mixOmics package (v6.10.9).

# 4.6. Data and Software Availability

The original contributions of this study are provided in the Supplementary Material.

# 4.7. Statistical Analysis

All the statistical analyses were performed by using R software (version 4.3.0). All the results are presented as the mean  $\pm$  standard deviation (SD), and a *p* value < 0.05 was considered statistically significant in our study.

**Supplementary Materials:** The following supporting information can be downloaded at https: //www.mdpi.com/article/10.3390/ijms26115146/s1.

**Author Contributions:** Conceptualization, L.H. and B.C.; methodology, B.C. and F.Z.; software, B.C.; validation, J.X. and B.C.; formal analysis, S.C.; investigation, X.Y.; resources, L.H.; data curation, B.C.; writing—original draft preparation, B.C. and L.H.; writing—review and editing, F.Z.; visualization, B.C.; supervision, F.Z and B.C.; project administration, F.Z.; funding acquisition, F.Z. and B.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key Research and Development Program of China (2022YFC2503100), National Natural Science Foundation of China (82304265, 82361138566), and Key Laboratory of Environmental Pollution and Disease Control of the Ministry of Education (Guizhou Medical University)/High Level School of Public Health of Guizhou Medical University Building Open Project Funding Project and Joint Key Laboratory of Endemic Diseases Open Project Funding Project (GMU-2024-HJZ-LH08).

**Institutional Review Board Statement:** This study was conducted in accordance with the Declaration of Helsinki, and it was approved by the Institutional Review Board (or Ethics Committee) of Xi'an Jiaotong University (protocol code LLSBPJ-2023-296 and 28.2.2023 date of approval).

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

**Data Availability Statement:** All the transcriptomic (RNA-seq) and proteomic (DIA-MS) data generated in this study have been deposited in the figshare repository under accession https://doi.org/10.6084/m9.figshare.29060792.v1.

Conflicts of Interest: The authors declare no conflicts of interest.

# References

- 1. Xu, J.; Wang, J.; Zhao, H. The Prevalence of Kashin-Beck Disease in China: A Systematic Review and Meta-Analysis. *Biol. Trace Elem. Res.* 2023, 201, 3175–3184. [CrossRef]
- Xie, D.; Liao, Y.; Yue, J.; Zhang, C.; Wang, Y.; Deng, C.; Chen, L. Effects of five types of selenium supplementation for treatment of Kashin-Beck disease in children: A systematic review and network meta-analysis. *BMJ Open* 2018, *8*, e017883. [CrossRef] [PubMed]
- Xiong, G. Diagnostic, clinical and radiological characteristics of Kashin-Beck disease in Shaanxi Province, PR China. *Int. Orthop.* 2001, 25, 147–150. [CrossRef]
- Cheng, B.; Liang, C.; Yang, X.; Li, P.; Liu, L.; Cheng, S.; Jia, Y.; Zhang, L.; Ma, M.; Qi, X.; et al. Genetic association scan of 32 osteoarthritis susceptibility genes identified TP63 associated with an endemic osteoarthritis, Kashin-Beck disease. *Bone* 2021, 150, 115997. [CrossRef]
- Ning, Y.; Hu, M.; Chen, S.; Zhang, F.; Yang, X.; Zhang, Q.; Gong, Y.; Huang, R.; Liu, Y.; Chen, F.; et al. Investigation of selenium nutritional status and dietary pattern among children in Kashin-Beck disease endemic areas in Shaanxi Province, China using duplicate portion sampling method. *Environ. Int.* 2022, 164, 107255. [CrossRef]
- 6. Yu, F.; Duan, C. RNA-seq analysis reveals different gene ontologies and pathways in rheumatoid arthritis and Kashin-Beck disease. *Int. J. Rheum. Dis.* **2018**, *21*, 1686–1694. [CrossRef] [PubMed]
- Wu, W.; He, A.; Wen, Y.; Xiao, X.; Hao, J.; Zhang, F.; Guo, X. Comparison of microRNA expression profiles of Kashin-Beck disease, osteoarthritis and rheumatoid arthritis. *Sci. Rep.* 2017, 7, 540. [CrossRef]

- Wu, C.; Liu, H.; Zhang, F.; Shao, W.; Yang, L.; Ning, Y.; Wang, S.; Zhao, G.; Lee, B.J.; Lammi, M. Long noncoding RNA expression profile reveals lncRNAs signature associated with extracellular matrix degradation in kashin-beck disease. *Sci. Rep.* 2017, *7*, 17553. [CrossRef] [PubMed]
- Duan, C.; Guo, X.; Zhang, X.D.; Yu, H.J.; Yan, H.; Gao, Y.; Ma, W.J.; Gao, Z.Q.; Xu, P.; Lammi, M. Comparative analysis of gene expression profiles between primary knee osteoarthritis and an osteoarthritis endemic to Northwestern China, Kashin-Beck disease. *Arthritis Rheum.* 2010, 62, 771–780. [CrossRef]
- 10. Hon, C.C.; Ramilowski, J.A.; Harshbarger, J.; Bertin, N.; Rackham, O.J.; Gough, J.; Denisenko, E.; Schmeier, S.; Poulsen, T.M.; Severin, J.; et al. An atlas of human long non-coding RNAs with accurate 5' ends. *Nature* **2017**, *543*, 199–204. [CrossRef]
- 11. Dai, Y.; Jian, C.; Wang, X.; Dai, X. Comprehensive expression profiles of mRNAs, lncRNAs and miRNAs in Kashin-Beck disease identified by RNA-sequencing. *Mol. Omics* **2022**, *18*, 154–166. [CrossRef] [PubMed]
- 12. Stahl, D.C.; Swiderek, K.M.; Davis, M.T.; Lee, T.D. Data-controlled automation of liquid chromatography/tandem mass spectrometry analysis of peptide mixtures. J. Am. Soc. Mass Spectrom. **1996**, 7, 532–540. [CrossRef] [PubMed]
- 13. Venable, J.D.; Dong, M.Q.; Wohlschlegel, J.; Dillin, A.; Yates, J.R. Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat. Methods* **2004**, *1*, 39–45. [CrossRef]
- Guo, T.; Kouvonen, P.; Koh, C.C.; Gillet, L.C.; Wolski, W.E.; Röst, H.L.; Rosenberger, G.; Collins, B.C.; Blum, L.C.; Gillessen, S.; et al. Rapid mass spectrometric conversion of tissue biopsy samples into permanent quantitative digital proteome maps. *Nat. Med.* 2015, 21, 407–413. [CrossRef]
- 15. Bundgaard, L.; Åhrman, E.; Malmström, J.; Auf dem Keller, U.; Walters, M.; Jacobsen, S. Effective protein extraction combined with data independent acquisition analysis reveals a comprehensive and quantifiable insight into the proteomes of articular cartilage and subchondral bone. *Osteoarthr. Cartil.* **2022**, *30*, 137–146. [CrossRef]
- KA, L.C.; Boitard, S.; Besse, P. Sparse PLS discriminant analysis: Biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinform.* 2011, 12, 253.
- Deng, H.; Lin, X.; Xiang, R.; Bao, M.; Qiao, L.; Liu, H.; He, H.; Wen, X.; Han, J. Low selenium and T-2 toxin may be involved in the pathogenesis of Kashin-Beck disease by affecting AMPK/mTOR/ULK1 pathway mediated autophagy. *Ecotoxicol. Environ. Saf.* 2024, 279, 116503. [CrossRef]
- Liu, R.; Xiao, Y.; Huang, S.; Wu, H.; Dong, J.; Zeng, S.; Li, Y.; Ye, J.; Wu, W.; Wang, M.; et al. LncRNA XIST inhibits mitophagy and increases mitochondrial dysfunction by promoting BNIP3 promoter methylation to facilitate the progression of KBD. *Mol. Immunol.* 2025, 182, 62–75. [CrossRef]
- 19. Wang, X.; Ning, Y.; Li, C.; Gong, Y.; Huang, R.; Hu, M.; Poulet, B.; Xu, K.; Zhao, G.; Zhou, R.; et al. Alterations in the gut microbiota and metabolite profiles of patients with Kashin-Beck disease, an endemic osteoarthritis in China. *Cell Death Dis.* **2021**, *12*, 1015. [CrossRef]
- 20. Buckwalter, J.A.; Anderson, D.D.; Brown, T.D.; Tochigi, Y.; Martin, J.A. The Roles of Mechanical Stresses in the Pathogenesis of Osteoarthritis: Implications for Treatment of Joint Injuries. *Cartilage* **2013**, *4*, 286–294. [CrossRef]
- 21. Liu, J.; Wang, L.; Guo, X.; Pang, Q.; Wu, S.; Wu, C.; Xu, P.; Bai, Y. The role of mitochondria in T-2 toxin-induced human chondrocytes apoptosis. *PLoS ONE* **2014**, *9*, e108394. [CrossRef] [PubMed]
- 22. Xu, J.; Fang, X.; Qin, L.; Wu, Q.; Zhan, X. LncRNA PVT1 regulates biological function of osteoarthritis cells by regulating miR-497/AKT3 axis. *Medicine* 2022, 101, e31725. [CrossRef] [PubMed]
- 23. Bartel, D.P. MicroRNAs: Target recognition and regulatory functions. Cell 2009, 136, 215–233. [CrossRef]
- 24. Wang, W.; Yu, Y.; Hao, J.; Wen, Y.; Han, J.; Hou, W.; Liu, R.; Zhao, B.; He, A.; Li, P.; et al. Genome-wide DNA methylation profiling of articular cartilage reveals significant epigenetic alterations in Kashin-Beck disease and osteoarthritis. *Osteoarthr. Cartil.* 2017, 25, 2127–2133. [CrossRef]
- 25. Kitahara, K.; Ebata, T.; Liyile, C.; Nishida, Y.; Ogawa, Y.; Tokuhiro, T.; Shiota, J.; Nagano, T.; Takasuka, T.E.; Endo, T.; et al. Chondroprotective functions of neutrophil-derived extracellular vesicles by promoting the production of secreted frizzled-related protein 5 in cartilage. *Cell Commun. Signal.* **2024**, *22*, 569. [CrossRef]
- 26. Gao, J.; Ren, J.; Ye, H.; Chu, W.; Ding, X.; Ding, L.; Fu, Y. Thymosin beta 10 loaded ZIF-8/sericin hydrogel promoting angiogenesis and osteogenesis for bone regeneration. *Int. J. Biol. Macromol.* **2024**, *267 Pt 1*, 131562. [CrossRef] [PubMed]
- 27. Park, S.J.; Kim, S.H.; Choi, H.S.; Rhee, Y.; Lim, S.K. Fibroblast growth factor 2-induced cytoplasmic asparaginyl-tRNA synthetase promotes survival of osteoblasts by regulating anti-apoptotic PI3K/Akt signaling. *Bone* **2009**, *45*, 994–1003. [CrossRef]
- Han, L.; Cheng, B.; Wei, W.; Liu, L.; Cheng, S.; Liu, H.; Jia, Y.; Wen, Y.; Zhang, F. Whole-Transcriptome Sequencing of Knee Joint Cartilage from Kashin-Beck Disease and Osteoarthritis Patients. *Int. J. Mol. Sci.* 2024, 25, 4348. [CrossRef]
- 29. Binvignat, M.; Emond, P.; Mifsud, F.; Miao, B.; Courties, A.; Lefèvre, A.; Maheu, E.; Crema, M.D.; Klatzmann, D.; Kloppenburg, M.; et al. Serum tryptophan metabolites are associated with erosive hand osteoarthritis and pain: Results from the DIGICOD cohort. *Osteoarthr. Cartil.* **2023**, *31*, 1132–1143. [CrossRef]
- 30. Zheng, H.; Xie, W. The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.* **2019**, 20, 535–550. [CrossRef]

- Carpenter, S.; Ricci, E.P.; Mercier, B.C.; Moore, M.J.; Fitzgerald, K.A. Post-transcriptional regulation of gene expression in innate immunity. *Nat. Rev. Immunol.* 2014, 14, 361–376. [CrossRef] [PubMed]
- 32. To, K.; Fei, L.; Pett, J.P.; Roberts, K.; Blain, R.; Polański, K.; Li, T.; Yayon, N.; He, P.; Xu, C.; et al. A multi-omic atlas of human embryonic skeletal development. *Nature* 2024, 635, 657–667. [CrossRef] [PubMed]
- 33. Rohart, F.; Gautier, B.; Singh, A.; KA, L.C. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* **2017**, *13*, e1005752. [CrossRef]
- 34. Hu, Z.; Zhao, J.; Liu, X.; Li, Y.; Jiang, H.; Fang, W.; Long, X. Glycyrrhizin regulates antioxidation through Nrf2 signaling pathway in rat temporomandibular joint osteoarthritis. *J. Oral Rehabil.* **2024**, *51*, 611–622. [CrossRef]
- Yang, M.; Peng, Y.; Liu, W.; Zhou, M.; Meng, Q.; Yuan, C. Sirtuin 2 expression suppresses oxidative stress and senescence of nucleus pulposus cells through inhibition of the p53/p21 pathway. *Biochem. Biophys. Res. Commun.* 2019, 513, 616–622. [CrossRef]
- 36. Morikawa, D.; Nojiri, H.; Saita, Y.; Kobayashi, K.; Watanabe, K.; Ozawa, Y.; Koike, M.; Asou, Y.; Takaku, T.; Kaneko, K.; et al. Cytoplasmic reactive oxygen species and SOD1 regulate bone mass during mechanical unloading. *J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res.* 2013, 28, 2368–2380. [CrossRef]
- 37. Sahasrabudhe, S.A.; Terluk, M.R.; Kartha, R.V. N-acetylcysteine Pharmacology and Applications in Rare Diseases-Repurposing an Old Antioxidant. *Antioxidants* **2023**, *12*, 1316. [CrossRef]
- 38. Liu, L.; Luo, P.; Wen, P.; Xu, P. Effects of selenium and iodine on Kashin-Beck disease: An updated review. *Front. Nutr.* **2024**, *11*, 1402559. [CrossRef] [PubMed]
- Liu, L.; Bai, L.; Wu, Y.; Zhang, Y.; Wang, C.; Li, S.; Yuan, Y.; Lv, X.; Wang, H.; Qin, Y.; et al. Exploring the role of gut microbiota in the pathogenesis of Kashin-Beck Disease: A Focus on selenium deficiency and T-2 toxin exposure. *Med. Hypotheses* 2025, 197, 111606. [CrossRef]
- 40. WS/T 207-2010; Diagnosis of Kashin-Beck Disease. National Health and Family Planning Commission of People's Republic of China: Beijing, China, 2010.
- 41. Young, M.D.; Wakefield, M.J.; Smyth, G.K.; Oshlack, A. Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.* **2010**, *11*, R14. [CrossRef]
- 42. Mao, X.; Cai, T.; Olyarchuk, J.G.; Wei, L. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* **2005**, *21*, 3787–3793. [CrossRef]
- 43. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. 2011, 17, 10–12. [CrossRef]
- 44. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. Nat. Methods 2012, 9, 357–359. [CrossRef] [PubMed]
- Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* 2015, *12*, 357–360. [CrossRef] [PubMed]
- 46. Pertea, M.; Pertea, G.M.; Antonescu, C.M.; Chang, T.C.; Mendell, J.T.; Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **2015**, *33*, 290–295. [CrossRef]
- 47. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [CrossRef]
- 48. Kong, L.; Zhang, Y.; Ye, Z.Q.; Liu, X.Q.; Zhao, S.Q.; Wei, L.; Gao, G. CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* 2007, *35*, W345–9. [CrossRef]
- 49. Sun, L.; Luo, H.; Bu, D.; Zhao, G.; Yu, K.; Zhang, C.; Liu, Y.; Chen, R.; Zhao, Y. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* **2013**, *41*, e166. [CrossRef]
- 50. Trapnell, C.; Williams, B.A.; Pertea, G.; Mortazavi, A.; Kwan, G.; van Baren, M.J.; Salzberg, S.L.; Wold, B.J.; Pachter, L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **2010**, *28*, 511–515. [CrossRef]
- 51. Conesa, A.; Götz, S.; García-Gómez, J.M.; Terol, J.; Talón, M.; Robles, M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 2005, 21, 3674–3676. [CrossRef]
- 52. Zhang, G.B.; Li, H.Y.; Yu, W.J.; Ying, Y.Z.; Zheng, D.; Zhang, X.K.; Wang, Y.G.; Shi, G.Z.; Huang, H.W. Occurrence and risk factors for post-stroke delirium: A systematic review and meta-analysis. *Asian J. Psychiatry* **2024**, *99*, 104132. [CrossRef] [PubMed]
- 53. Singh, A.; Shannon, C.P.; Gautier, B.; Rohart, F.; Vacher, M.; Tebbutt, S.J.; KA, L.C. DIABLO: An integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* **2019**, *35*, 3055–3062. [CrossRef] [PubMed]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.